

Random Amplification and Pyrosequencing for Identification of Novel Viral Genome Sequences

Jun Hang,^{1,*} Brett M. Forshey,² Tadeusz J. Kochel,^{2,3} Tao Li,¹ Victor Fiestas Solórzano,⁴ Eric S. Halsey,² and Robert A. Kuschner¹

¹Viral Diseases Branch, Walter Reed Army Institute of Research, Silver Spring, Maryland 20910, USA; ²U.S. Naval Medical Research Unit Six, Iquitos and Lima, Perú; ³U.S. Naval Medical Research Center, Silver Spring, Maryland 20910, USA; and ⁴Centro Nacional de Salud Pública, Instituto Nacional de Salud, Lima, Perú

ssRNA viruses have high levels of genomic divergence, which can lead to difficulty in genomic characterization of new viruses using traditional PCR amplification and sequencing methods. In this study, random reverse transcription, anchored random PCR amplification, and high-throughput pyrosequencing were used to identify orthobunyavirus sequences from total RNA extracted from viral cultures of acute febrile illness specimens. Draft genome sequence for the orthobunyavirus L segment was assembled and sequentially extended using de novo assembly contigs from pyrosequencing reads and orthobunyavirus sequences in GenBank as guidance. Accuracy and continuous coverage were achieved by mapping all reads to the L segment draft sequence. Subsequently, RT-PCR and Sanger sequencing were used to complete the genome sequence. The complete L segment was found to be 6936 bases in length, encoding a 2248-aa putative RNA polymerase. The identified L segment was distinct from previously published South American orthobunyaviruses, sharing 63% and 54% identity at the nucleotide and amino acid level, respectively, with the complete Oropouche virus L segment and 73% and 81% identity at the nucleotide and amino acid level, respectively, with a partial Caraparu virus L segment. The result demonstrated the effectiveness of a sequence-independent amplification and next-generation sequencing approach for obtaining complete viral genomes from total nucleic acid extracts and its use in pathogen discovery.

KEY WORDS: orthobunyavirus, next-generation DNA sequencing, microbial, pathogen discovery

INTRODUCTION

Advanced, massively parallel DNA sequencing—so-called next-generation sequencing technology—has been changing the landscape of microbial genomics research.^{1,2} It has

been applied extensively in studies otherwise too unmanageable or expensive to conduct, such as sequencing unculturable microbes and sequencing a mixed population without isolation and cloning. Furthermore, characterization of novel viruses can be facilitated by using next-generation sequencing technology, thereby overcoming the challenges posed by the high degree of genetic diversity observed across most virus families and the lack of available, full-length sequences for comparison. In light of the rapidly increasing capacity and cost-effectiveness of the technology, there is a need to improve upstream sample preparation techniques and to facilitate downstream data analysis.³

Only a limited number of studies have been conducted on the South American orthobunyaviruses (family: *Bunyaviridae*; genus: *Orthobunyavirus*), and only a handful of full-length sequences are available for analysis, making this genus a perfect candidate for exploration with a next-generation sequencing approach. Orthobunyaviruses are arthropod-borne pathogens with trisegmented, negative-sense RNA genomes and a worldwide distribution, and they can cause serious illness in humans and animals.^{4,5}

*ADDRESS CORRESPONDENCE TO: Jun Hang, Viral Diseases Branch, Walter Reed Army Institute of Research, 503 Robert Grant Ave., Silver Spring, MD 20910, USA (E-mail: jun.hang.ctr@us.army.mil).

DISCLAIMERS: The opinions expressed in this work are those of the authors and do not reflect the official policy or position of the Department of the Navy, Department of the Army, Department of the Air Force, DoD, or U.S. government. We declare that no conflict of interest exists.

Copyright Statement. Some authors are military service members (or employees of the U.S. government). This work was prepared as part of their official duties. Title 17 U.S.C. §105 provides that “Copyright protection under this title is not available for any work of the United States Government”. Title 17 U.S.C. §101 defines a U.S. government work as a work prepared by a military service member or employee of the U.S. government as part of that person’s official duties.

doi: 10.7171/jbt.12-2301-001

Human infection often causes fever, headache, myalgia, nausea, vomiting, and weakness but can also result in encephalitis or multiorgan dysfunction. However, molecular biology and genetic studies about South American orthobunyaviruses to this point have been limited.⁶ As a consequence, PCR-based studies often encounter difficulties, owing to a lack of sufficient genome sequences available for primer design and genomic comparisons. To address this deficiency and enhance our understanding of orthobunyavirus transmission patterns in the region, comprehensive genome studies about a variety of orthobunyaviruses are needed.

Next-generation DNA sequencing has been used successfully in the study of various viral genomes, including exploration of a number of orthobunyaviruses.⁶ Herein, we describe the development of a method for efficient sequencing of unknown viral isolates from tissue culture and apply these techniques to samples from Peru, which were tentatively classified as orthobunyaviruses by immunoassay but were refractory to RT-PCR amplification by primers designed based on known sequences. The procedure entailed anchored, random reverse transcription and PCR, pyrosequencing, and data mining and analysis using several tools. The developed method will be applied to the comprehensive study of a large collection of orthobunyaviruses isolated from clinical specimens and mosquito samples.⁷

MATERIALS AND METHODS

Collection of Clinical Specimens and Viral Culture

Acute-phase serum samples were collected from a 14-year-old female (sample code IQE7620) and a 46-year-old female (IQE7743), who reported to clinics outside of Iquitos, Peru, with acute, undifferentiated febrile illnesses in April and July 2008, respectively. Acute-phase serum samples were inoculated onto C6/36 (*Aedes albopictus*) cells and Vero (African green monkey) cells for virus isolation, as described previously.⁷

Immunoassay, Total RNA Purification, and Amplifications

Upon observation of cytopathic effect, tissue-culture wells were probed for viral antigens by a standard, indirect immunofluorescence assay using hyperimmune mouse ascitic fluid, produced against a variety of arthropod-borne viruses, including orthobunyaviruses from the Group C complex [Caraparu virus (CARV) and Murutucu virus (MURV)]. Following detection of viral antigen, tissue-culture supernatants were collected, and RNA was extracted using the QIAamp viral RNA extraction kit (Qiagen, Valencia, CA, USA). To detect Group C orthobunyavirus RNA, RT-PCR was attempted on cultured viruses collected from Vero cells using primers and

conditions described previously for detecting S segment (primer pair BUN-F/BUN-R) and M segment (primer pairs M14/M619R and BUN-GnF/BUN-GnR) sequences.^{8,9}

Random RT-PCR and Amplicon Pyrosequencing

The total RNA extract was subjected to reverse transcription and PCR amplification. The method was based on the published methods and modified for improvement.^{10–12} A commercial product, Ovation RNA-Seq system (NuGEN Technologies, San Carlos, CA, USA), was used following the manufacturer's instruction for comparison with the method. In the modified, anchored, random RT-PCR method, 10–100 ng total RNA in 11.5 μ l 10 mM Tris-HCl, pH 7.5 or 8.0, was mixed with 1 μ l 10 mM dNTP mix and 0.5 μ l 100 μ M primer P17N8 (sequence shown in Table 1), incubated in a GeneAmp PCR System 9700 (Applied Biosystems, Foster City, CA, USA) at 65°C for 5 min, and then cooled down to 4°C. Random RT-PCR for cDNA synthesis was conducted by adding 4 μ l 5 \times First Strand synthesis buffer, 1 μ l 0.1 M DTT, 1 μ l RNaseOUT, and 1 μ l SuperScript III RT (Invitrogen, Carlsbad, CA, USA) and incubation at 25°C for 5 min, 50°C for 60 min, and 75°C for 15 min. One unit of RNase H (Invitrogen) was added into the reaction mixture and incubated for 20 min at 37°C to digest RNA. dsDNA was synthesized by PCR amplification in a 50- μ l reaction containing 2 μ l First Strand synthesis reaction, 0.5 μ l 10 μ M primer P17N8, 4.5 μ l 10 μ M primer P21 (sequence shown in Table 1), 1 μ l 10 mM dNTPs, 1 μ l Platinum Taq DNA Polymerase High Fidelity (5 U/ μ l), 5 μ l 10 \times PCR buffer without magnesium, and 1.5 μ l 50 mM MgCl₂. The thermal cycling program for the PCR reaction was 94°C for 2 min, 25°C for 30 s, and 72°C for 1 min and then 20–30 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 1 min, followed by 72°C for 5 min and 4°C after the reaction. The amplification products were examined by agarose gel electrophoresis and then purified with the PurElute GX DNA Gel Extraction and Cleanup Kit (EdgeBio, Gaithersburg, MD, USA).

A RL amplicon pyrosequencing library for the purified, anchored, random RT-PCR products was prepared by using the genome sequencer (GS) FLX Titanium Rapid Library Preparation Kit (Roche 454 Life Sciences, Branford, CT, USA). After purification, the DNA size distribution of the RL library was examined by PCR assay using primers 454 LinkerA and 454 LinkerB (sequences shown in Table 1). There was no size selection of the library performed, unless there was a large quantity of small fragments (i.e., ≤ 100 bp) in the PCR assay result. In size selection to remove the short fragments, the RL library was resolved on a 2% agarose gel; the smeared band with the size ≥ 150 bp was collected and subjected to DNA gel

TABLE 1

Primers Used in the Study for Reverse Transcription, PCR Amplifications, and Sanger Sequencing

Primer Name	Primer Description	Nucleotide Sequence
P17N8	Random RT, anchored PCR ¹²	GTTTCCCAGTAGGTCTC NNNNNNNNN ^a
P21	Anchored PCR ¹²	CGCCGTTTCCCAGTAGGTCTC
454LinkerA		GCCTCCCTCGCGCCATCAG
454LinkerB	Primer pair for PCR amplification of RL library	GCCTTGCCAGCCCCGCTCAG
1LP5END25fwd	P17 plus 8-base <u>consensus sequence</u>	GTTTCCCAGTAGGTCTCAGTAGTGT ^b
1LP5END22fwd	P17 plus 5-base <u>consensus sequence</u>	GTTTCCCAGTAGGTCTCAGTAG ^b
1LP261rev	Reverse primer for 5' end	GTTCTTAGATCCAGATCAGGTGCC
1LP6636fwd	Forward primer for 3' end	GCCAAGGAAACCACTGGGATCAAA
1LP592fwd		TTAGTGGCTCATGGTGA
1LP1819rev	Primer pair for a gap between de novo contigs	AACCACCCTTGACCTGTATGTCCT
1LP4858fwd		CCCTGCATTAGTGCTGAGGGCT
1LP5146rev	Primer pair for a gap between de novo contigs	TGCAGTGTATGCCTTGGACGGC
1LP5END25fwd		GTTTCCCAGTAGGTCTCAGTAGTGT
1LP261rev	Primer pair for determining 5' terminal sequence	GTTCTTAGATCCAGATCAGGTGCC
1LP6332fwd		TTGCAAGATGCACTGCCGAGA
1LP5END25fwd	Primer pair for determining 3' terminal sequence	GTTTCCCAGTAGGTCTCAGTAGTGT

RL, Rapid ligation. ^aRandom octamer oligo at 3'-terminus. ^bUnderlined, partial consensus 5'-terminal orthobunyavirus L segment nucleotide sequence.⁹

extraction. The RL library copy number concentration was determined using a RL standard and a QuantiFluor-P fluorometer (Promega, Madison, WI, USA). Subsequent emulsion PCR and pyrosequencing on the GS FLX system (Roche 454 Life Sciences) were performed by following standard protocols, except that PCR primer concentration was reduced by 75% in amplicon emulsion.

Pyrosequencing Data Analysis and Sequence Assembling and Finishing

The sequence reads produced by Roche GS FLX Titanium pyrosequencing were assembled to generate consensus contigs using Roche GS FLX software GSAssembler version 2.3. All individual reads were subjected to direct, discontinuous Megablast (word length 11; <http://blast.ncbi.nlm.nih.gov/>) against GenBank's nucleotide collection (nr/nt) to identify viral genome(s) with the highest similarity. A more-sensitive search was also performed, using the word length 7 in mpiBLAST installed on Department of Defense (DoD) high-performance computing (HPC) Linux clusters.¹³ GS de novo assembly contigs were aligned with an identified viral genome sequence to determine the order and orientation of the contig sequences. Based on the mapping results, a concatenated sequence corresponding to a partial genome sequence was compiled with a set of selected contigs and iteratively extended semimanually using an approach similar to a recently published, automated algorithm.¹⁴ Individual reads overlapping the ends of the concatenated sequence and/or other contigs were identified through alignment using the BLAST tool. The process was

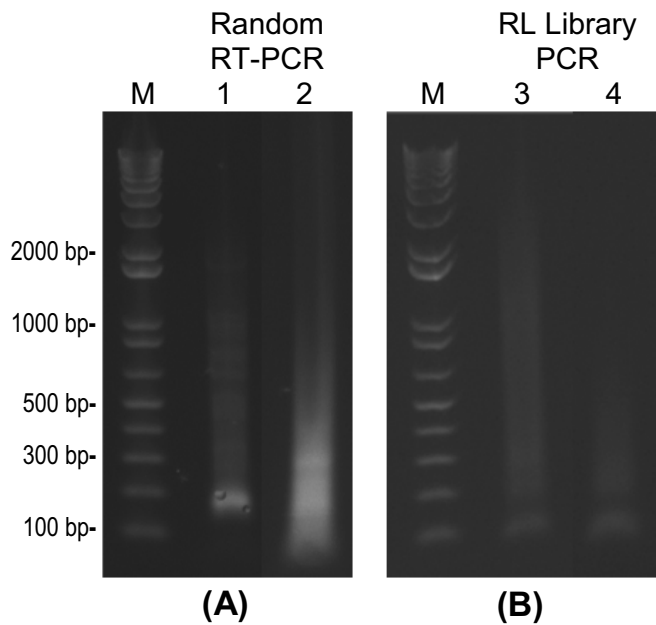
repeated until the assembled and extended sequence reached the maximal length on both ends.

Primer sets were designed based on the genome draft sequence (Table 1). The chimeric primer corresponding to oligo P21, plus the orthobunyavirus consensus terminal sequence, was used to amplify the sequences at both ends.⁹ Additional primer pairs were used to amplify the intermediate regions corresponding to the junctions or gaps between de novo assembly contigs, which were used in compiling the draft genome sequence. PCR products were purified and sequenced by the Sanger method using Big-Dye Terminator v3.1 (Applied Biosystems). Sequencing reactions were purified by alcohol precipitation, resuspended in 10 µl Hi-Di formamide, and analyzed on a 3130 xl Plus DNA analyzer (Applied Biosystems).

RESULTS AND DISCUSSION

Viral Culture, Immunoassay, and Molecular Analyses

Inoculation of the clinical specimens led to the development of the cytopathic effect on C6/36 cells and Vero cells. Viral cultures of both isolates (IQE7620 and IQE7743) were reactive with CARV and MURV antibodies in standard, indirect immunofluorescence assays, indicating the presence of orthobunyavirus in the clinical isolates. Amplification of viral RNA was attempted using primer pairs described previously targeting S and M segment sequences; however, no cDNA was generated, despite repeated attempts under varied RT-PCR and PCR conditions. The inconclusive results suggested the probable presence of an

**FIGURE 1**

Random RT-PCR amplification of total RNA extract and PCR amplification of RL library for quality assessment. (A) Total RNA extract from IQE7620 viral culture supernatant was amplified using an anchored, random amplification method (Lane 1) or NuGEN Technologies' RNA-Seq system (Lane 2). (B) RL libraries were amplified using 454 LinkerA and 454 LinkerB as primers. The templates for PCR were anchored, random amplification products (Lane 3) or NuGEN Technologies' RNA-Seq amplification products (Lane 4). One percent E-gel and 1 kb plus DNA ladder (Invitrogen) were used for the agarose gel electrophoresis analysis.

orthobunyavirus with a novel genome sequence sufficiently different from published sequences.

Viral Genome Sequence Identification by Pyrosequencing

A new approach was attempted to overcome the technical difficulty in identification of viral genome RNA. The approach was comprised of two major experimental processes: (1) First Strand cDNA synthesis by random reverse transcription, followed by random PCR amplification for synthesis of dsDNA amplicons (random RT-PCR), and (2) ultrahigh-throughput pyrosequencing of the amplicons. Two random RT-PCR methods were tested for comparison. The modified, anchored, random RT-PCR and NuGEN Technologies' Ovation RNA-Seq system were productive in amplification, but larger fragment sizes were generated using the anchored, random RT-PCR approach as compared with NuGEN Technologies' method (Fig. 1). The result was not surprising, as NuGEN Technologies' Ovation RNA-Seq system was developed mainly for downstream application on the Illumina next-generation sequencing platform, which optimally requires small DNA

fragments for its short read sequencing. Anchored RT-PCR generated amplicons with a broad size distribution from 200 bp to >1 kb. No outstanding DNA bands were observed, suggesting an unbiased randomness of the method. The relatively larger sizes of the amplicons are suited for efficient and productive pyrosequencing on the Roche GS FLX Titanium system, which typically generates read lengths of ~350 bp. Small RNA segments for many RNA virus genome segments are barely 1 kb or shorter, often resulting in small fragments in random RT-PCR amplification. Therefore, in most cases, we chose to bypass a size-selection or small-fragment removal step to avoid biased selection and loss of sequences from small genome segments or in particular, structural regions that are less-efficiently amplified. However, in emulsion PCR, the small fragments in the amplicon library were amplified on the sequencing beads much more efficiently than the large fragments, resulting in highly variable intensity from well to well during pyrosequencing, resulting in poor raw images and a low quality of sequencing reads. Among several feasible approaches, we chose to suppress emulsion PCR amplification for small fragments by reducing the primer concentration by 75%. As a result, from a pyrosequencing run having multiple samples on an eight-region PicoTiter Plate (PTP), we obtained ~150,000 reads totaling 30 Mb of sequence for the isolate IQE7620 from the area equivalent to 15.6% of a PTP.

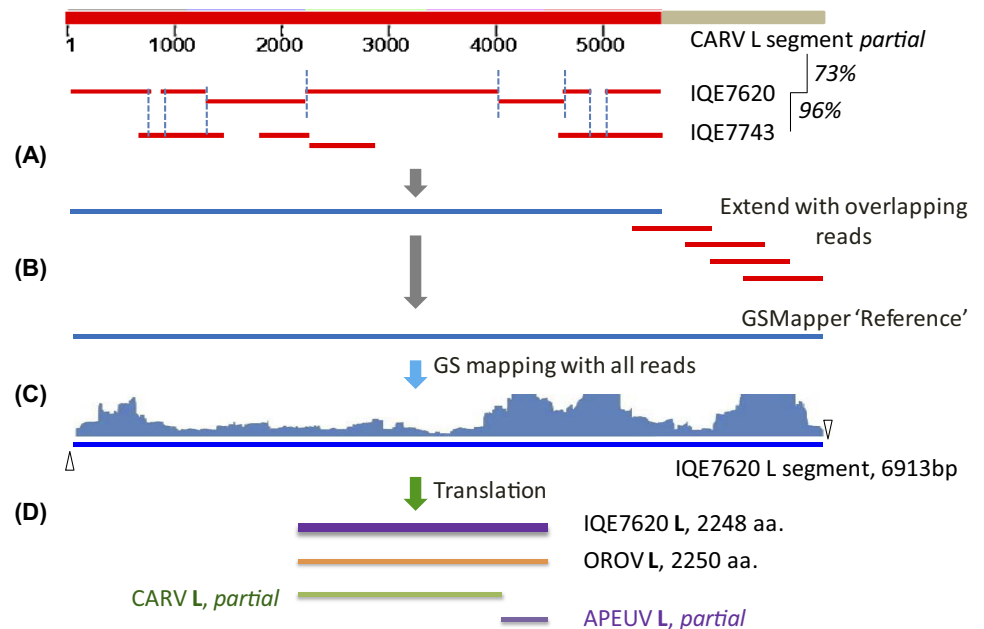
To determine if an orthobunyavirus sequence was present, we used an unbiased approach, and all individual reads were first compared against nucleotide collection (nr/nt) in GenBank using mpiBLAST, as described above. The BLAST hits were then examined for viral sequences using a Perl script that was developed in-house. Thousands of viral sequence hits were identified, mostly belonging to the orthobunyavirus family, with the highest sequence similarity to CARV.¹⁵ Approximately 65% identity was observed with the CARV L segment—too low for the CARV genome sequence to be used as a reference for mapping the raw reads with software GSMapper to assemble the genome sequence of IQE7620. This result suggested the possibility of IQE7620 being a novel orthobunyavirus and the necessity of using a de novo approach to assemble its genome sequence.

Sequence Assembly and Finishing of the Novel Viral Genome

GS FLX system data analysis software, GSAssembler, was used to assemble raw reads without a reference genome sequence (i.e., de novo assembly). In total, 668 contigs (150 kb in total) were obtained from de novo assembly. The contigs were aligned with the CARV L segment sequence to identify homologous sequences (Fig. 2A). Seven

FIGURE 2

Sequence assembly process for generating the complete L segment sequence for orthobunyavirus IQE7620. (A) Pyrosequencing reads were de novo-assembled using GSAssembler. Contigs, matched with partial sequence to the CARV segment L, were aligned and assembled together. The yellow bar signifies the missed sequence on the 5' region. (B) The draft was sequentially extended by searching adjacent contigs or reads until it reached the maximal length. (C) The extended draft was used as a "reference" to map with all pyrosequencing reads using GSMapper. Plot for depth of coverage was shown. Δ , Sanger sequencing was used to determine terminal sequences, which were 12 and 11, respectively, at the two ends, in addition to the 6913-base sequence determined by pyrosequencing. (D) The finished, complete L segment sequence for isolate IQE7620 was translated to a single 2248-aa sequence and aligned to three RNA polymerase sequences of related orthobunyaviruses. OROV, Oropouche virus; APEUV, Apeu virus.



contigs (5378 bases in total) were identified, which covered 96.8% of the 5555-bp partial sequence of the CARV L segment. The size of a complete L segment for an orthobunyavirus is ~ 6900 bp.⁴ To identify additional sequences between contigs and at both ends, the National Center for Biotechnology Information (NCBI) Megablast alignment tool was used to search for individual reads with sequences overlapping the contigs and extended sequences adjacent to the contigs. In addition, two contigs for isolate IQE7743, which were 96% identical to IQE7620 in overlap regions, were also used to fill the gaps (Fig. 2A and B). The process was repeated until the assembly reached its maximal length, and no additional reads containing extra sequences were found (Fig. 2B). The resulting assembly was a 6917-bp draft genome sequence, the expected size of a full orthobunyavirus L segment. Finally, we used this assembled draft genome sequence as a tentative reference to map all pyrosequencing reads, $\sim 150,000$ reads in total, with GSMapper. A total of 5524 reads (1.46 Mb in total) was mapped, and a single contig of 6913 bp was obtained. The alignment coverage ranged from 23 to 1627, with an average coverage of 210 (Fig. 2C). The mapped reads constituted only 3.1% of all reads, suggesting a low percentage of IQE7620 viral nucleic acids in the total RNA isolated from the viral culture.

The resulting nucleotide sequence comprised an

open reading frame of 2248 aa. The deduced, putative protein had approximately the same size as the 2250-aa RNA polymerase protein of OROV, which was one of orthobunyavirus references chosen by NCBI.¹⁶ Interestingly, amino acid identity with OROV was low (54%). Although the generated L segment sequence already contained the complete coding region for a RNA polymerase, it did not have the terminal consensus sequence typical of orthobunyaviruses and was thus incomplete. Additionally, it was necessary to confirm the accuracy of the sequence because of the low identity to the reference and the novelty of the sequence. A set of primers was used to amplify sequences in the terminal, as well as several intermediate regions (Table 1). Instead of using a random RT-PCR product or the final random amplicon library, we used the reverse-transcription product (i.e., First Strand cDNA) as the template in PCR, in case there were amplification errors introduced in the anchored, random PCR. Purified PCR fragments were sequenced using the Sanger method. No mismatch was found. From the Sanger sequencing results, 12 and 11 bases, including terminal consensus sequences, were added to the 5' and 3' ends of the sequence, respectively. The final, complete L segment sequence for the IQE7620 genome was found to be 6936 bp, encoding a 2248-aa putative orthobunyavirus RNA polymerase.

TABLE 2

L Segment Sequence Similarity of ZUNV to the Orthobunyaviruses Bunyamwera Virus, La Crosse Virus, OROV, CARV, and APEUV

Virus	Description	Genome Sequence, Bases	Protein Sequence, Amino Acids	Sequence Identity (Sequence Similarity)	
				Nucleotide	Amino Acid
ZUNV	Isolate IQE7620, JN157805	6936, <i>complete</i>	2248, <i>complete cds</i>	–	–
Bunyamwera	NCBI RefSeq, NC_001925	6875, <i>complete</i>	2238, <i>complete cds</i>	61%	52% (70%)
La Crosse	NCBI RefSeq, NC_004108	6980, <i>complete</i>	2263, <i>complete cds</i>	61%	52% (71%)
OROV	NC_005776	6846, <i>complete</i>	2250, <i>complete cds</i>	63%	54% (73%)
CARV	Group C EF122411	5555, <i>partial</i>	1833, <i>partial cds</i>	73%	81% (92%)
APEUV	Group C FJ859039	1254, <i>partial</i>	418, <i>partial cds</i>	78%	90% (97%)

Sequence similarity shown in parentheses indicates percentage of chemically similar amino acids between protein sequences.

Because of the significant nucleotide divergence from other reported orthobunyavirus sequences, the sequence was provisionally designated as Zungarococha virus (ZUNV) L segment (in reference to the lake near where it was isolated) and submitted to GenBank (accession JN157805). Based on sequence identity, it was more closely related to OROV than two other NCBI reference orthobunyaviruses—La Crosse virus and Bunyamwera virus (Table 2). ZUNV was closely related to two Group C complex orthobunyaviruses—CARV and APEUV (Table 2).¹⁷ It should be noted that there was a paucity of South American orthobunyavirus sequences, particularly L segment sequences, available for comparison. Prior to this report, no full-length Group C orthobunyavirus L segments have been published.

The low-sequence identity of IQE7620 with available orthobunyavirus sequences in GenBank may help to explain the unsuccessful RT-PCR amplification using primers designed based on known sequences. The low purity of viral nucleic acids in the total RNA extract, as suggested by the low percentage of mapped reads, was also consistent with the unsuccessful outcome of searching orthobunyavirus by cDNA cloning and Sanger sequencing (data not shown). The results demonstrated the effectiveness of a massively parallel pyrosequencing approach for identification of unknown viruses. Notably, the entire high-quality genome sequence was obtained from total RNA extracts by our methodology. Traditionally, viral particle purification by ultracentrifugation or sample preprocessing is used for sample enrichment to facilitate the analysis. However, these pretreatments may introduce bias or cause a loss of low-abundant fractions, which may hinder or prevent the discovery of unexpected agents. In this study, a random RT-PCR method was used for amplification in an unbiased manner to target all sequences in the specimen. Interestingly, we found a large portion of sequences that belonged

to a few classes of nucleic acids, including mitochondrial DNA and ribosomal RNA from the host and DNA from mycoplasma, an organism common in clinical specimens and tissue culture (data not shown). Specific removal of these major contaminants likely will increase the efficiency of the method.

The methodology outlined here will be used in further aspects of the project, including completion of S and M segment sequences for the ZUNV isolate IQE7620. More isolates from different geographic sites and time periods will be sequenced for phylogenetic analyses and studies of orthobunyavirus evolution.^{5,9,18} The novel orthobunyavirus will be subjected to further virological investigations to reveal functional attributes, such as antigenic studies, and to define more precisely the genetic relationship with other orthobunyavirus group members.

CONCLUSIONS

With increasing productivity and decreasing costs for next-generation DNA sequencing technologies, it will be cost-effective and affordable to sequence viral specimens or culture isolates by using total nucleic acid extracts without tedious purification or enrichment of viruses. This study demonstrated the technical feasibility of using anchored, random RT-PCR and pyrosequencing to find novel virus and obtain complete genome sequences, despite the low abundance of the virus in the total nucleic acid extract. The method obviates virus purification and does not require known nucleotide sequences for amplification, thereby using an unbiased approach that harnesses the power of next-generation sequencing technology. The attributes of anchored, random RT-PCR and pyrosequencing make these techniques particularly useful for viral genome discovery and identification of unknown pathogens that elude detection through conventional means.

ACKNOWLEDGMENTS

The authors thank Drs. Leonard N. Binn, Robert J. Putnak, Stephen J. Thomas, Deidra J. Shuck-Lee, and Yong Luo for support and critical reading of the manuscript. We thank Ms. Heping Gong and Ms. Yu Yang for technical supports. We also thank Dr. Nela Zavaljevski and Mr. Valmik Desai of the DoD Biotechnology HPC Software Applications Institute for providing initial identification of the orthobunyavirus content in the samples based on the mpiBLAST search. The study was supported by DoD Global Emerging Infections Surveillance and Response System funding to Naval Medical Research Center Detachment (NMRCDD; work unit number: 847705.82000.25GB.B0016) and to Walter Reed Army Institute of Research, Division of Viral Diseases (10301_11_WR).

The clinical specimens used in this study were obtained under the terms of a human use protocol (NMRCDD.2000.0006), approved by the Naval Medical Research Center Institutional Review Board (Bethesda, MD, USA) and the Dirección General de Epidemiología of Peru in compliance with all U.S. federal regulations governing the protection of human subjects. Written, informed consent was obtained from the participants or from a parent or legal guardian.

REFERENCES

1. Petrosino JF, Highlander S, Luna RA, Gibbs RA, Versalovic J. Metagenomic pyrosequencing and microbial identification. *Clin Chem* 2009;55:856–866.
2. Metzker ML. Sequencing technologies—the next generation. *Nat Rev Genet* 2010;11:31–46.
3. Fuller CW, Middendorf LR, Benner SA, et al. The challenges of sequencing by synthesis. *Nat Biotechnol* 2009;27:1013–1023.
4. Elliott RM. Emerging viruses: the *Bunyaviridae*. *Mol Med* 1997;3:572–527.
5. Mores CN, Turell MJ, Dyer J, Rossi CA. Phylogenetic relationships among orthobunyaviruses isolated from mosquitoes captured in Peru. *Vector Borne Zoonotic Dis* 2009;9:25–32.
6. Savji, N., Palacios G, Travassos da Rosa A, et al. Genomic and phylogenetic characterization of Leanyer virus, a novel orthobunyavirus isolated in northern Australia. *J Gen Virol* 2011;92:1676–1687.
7. Forshey BM, Guevara C, Laguna-Torres VA, et al. Arboviral etiologies of acute febrile illnesses in Western South America, 2000–2007. *PLoS Negl Trop Dis* 2010;4:e787.
8. Bowen MD, Trappier SG, Sanchez AJ, et al. A reassortant bunyavirus isolated from acute hemorrhagic fever cases in Kenya and Somalia. *Virology* 2010;291:185–190.
9. Nunes MR, Travassos da Rosa AP, Weaver SC, Tesh RB, Vasconcelos PF. Molecular epidemiology of group C viruses (*Bunyaviridae*, orthobunyavirus) isolated in the Americas. *J Virol* 2005;79:10561–10570.
10. Djikeng A, Halpin R, Kuzmickas R, et al. Viral genome sequencing by random priming methods. *BMC Genomics* 2008;9:5.
11. Djikeng A, Spiro D. Advancing full length genome sequencing for human RNA viral pathogens. *Future Virol* 2009;4:47–53.
12. Briese T, Paweska JT, McMullan LK, et al. Genetic detection and characterization of Lujo virus, a new hemorrhagic fever-associated arenavirus from southern Africa. *PLoS Pathog* 2009;5:e1000455.
13. Darling AE, Carey L, Feng W. The design, implementation and evaluation of mpiBLAST. *Proceeding of the Fourth International Conference on Linux Clusters*. San Jose, CA, USA, 2003.
14. Tsai IJ, Otto TD, Berriman M. Improving draft assemblies by iterative mapping and assembly of short reads to eliminate gaps. *Genome Biol* 2010;11:R41.
15. de Brito Magalhães CL, Quinan BR, Novaes RF, et al. Carapar virus (group C orthobunyavirus): sequencing and phylogenetic analysis based on the conserved region 3 of the RNA polymerase gene. *Virus Genes* 2007;35:681–684.
16. Pruitt KD, Tatusova T, Klimke W, Maglott DR. NCBI reference sequences: current status, policy, and new initiatives. *Nucleic Acids Res* 2009;35:D32–D36.
17. de Brito Magalhães CL, Drumond BP, Novaes RF, et al. Identification of a phylogenetically distinct orthobunyavirus from group C. *Arch Virol* 2011;156:1173–1184.
18. Turell MJ, O'Guinn ML, Jones JW, et al. Isolation of viruses from mosquitoes (Diptera: *Culicidae*) collected in the Amazon Basin region of Peru. *J Med Entomol* 2005;42:891–898.